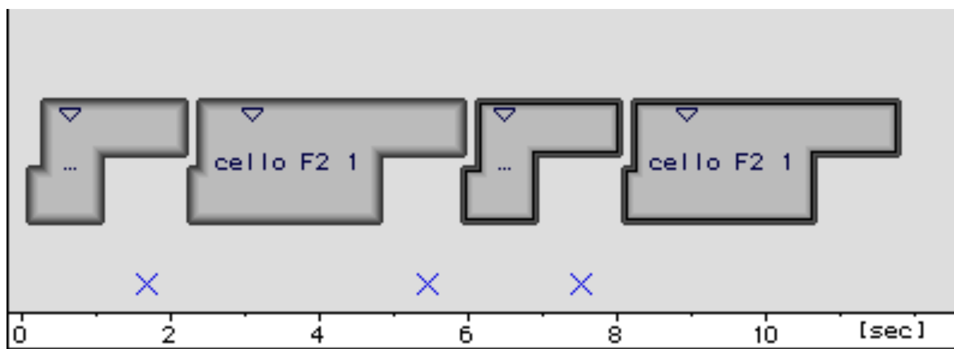
Ircam documentation

- Research reports
- Musical works
- **Software**

Diphone Studio (v. 2.8)

Obtention de Modèles de résonance



Première édition du manuel
novembre 1999

ircam
Centre Pompidou

 Il a été rédigé par Francesc Marti, d'après le document de Pierre-François Baisnée en novembre 1990.

Diphone
Conception : Xavier Rodet
Programmation Macintosh : Adrien Lefèvre
Programmation Unix : Thierry Galas, Philippe Depalle
Etudiants-stagiaires : Guillaume Bouriez, Marteen de Boer, Xavier Hosxe, Gilbert Nouno

SVP
Conception : Philippe Depalle
Programmation : Philippe Depalle, Gilles Poirot, Chris Rogers, Jean Carrive

Pm
Conception : Xavier Rodet et Guillermo Garcia
Programmation : Guillermo Garcia
Cette documentation correspond à la version 2.0 de Diphone (version Macintosh).

Première édition, novembre 1999.
Ce manuel correspond à ces versions de Diphone Studio
2.8 pour l'application Diphone, 1.8 pour l'application Addan, 1.0 pour l'application Resan.

Ircam Centre Pompidou
1, place Igor-Stravinsky
F-75004 Paris
Tel. 01 44 78 49 62
Fax 01 44 78 15 40
E-mail ircam-doc@ircam.fr

IRCAM Users' group

The use of this software and its documentation is restricted to members of the Ircam software users' group. For any supplementary information, contact:

Département de la Valorisation
Ircam
Place Stravinsky, F-75004 Paris

Tel. 01 44 78 49 62
Fax 01 44 78 15 40
E-mail: bousac@ircam.fr

Send comments or suggestions to the editor:
E-mail: bam@ircam.fr
Mail: Marc Battier,
Ircam, Département de la Valorisation
Place Stravinsky, F-75004 Paris

http://www.ircam.fr/forumnet



Pour voir la table des matières, cliquez
sur le bouton qui se trouve dans le bandeau
de la fenêtre Acrobat.
Pour se déplacer d’une page à l’autre, utilisez
les boutons de navigation d’Acrobat ou les
touches de flèches -> et <- de votre clavier.

Contenu

Introduction

Le travail de recherche autour des modèles de résonance a été mené à l'IRCAM en 1985 et 1986 par Yves Potard, Pierre-François Baisnée, et Jean-Baptiste Barrière. Des outils ont été développés en Lisp dans l'environnement CHANT-FORMES pour l'obtention et l'utilisation de modèles.

En 1989, une version en C de l'analyse comportant quelques améliorations a été écrite par P.-F. Baisnée lors d'un séjour aux Etats-Unis. Cette version a d'abord été adaptée en 1989 pour fonctionner dans le programme Macmix d'Adrian Freed (dans cette version, les modèles peuvent être présentés graphiquement à l'écran, et la synthèse a lieu en temps réel sur une machine développée par A. Freed et Marie-Dominique Baudot - « Reson8 »). Elle a ensuite été adaptée à l'IRCAM pour fonctionner dans un environnement UNIX ou MPW.

En 1999, Francesc Marti récupère l'ancien code et l'adapte pour fonctionner dans un environnement UNIX ou SGI. En plus, il s'est ajouté des petites modifications et améliorés.

Adrien Lefevre adapte ce code et il construit une version pour fonctionner dans le logiciel Diphone (dans un environnement Macintosh).

Le présent document se contente de donner un bref mode d'emploi de cette dernière version.

Les commandes et les machines utilisables

Quatre commandes ont été implémentées (exécutables UNIX ou SGI) : **modRes**, **modRes_scrip**, **reduce** et **mk_lisp_model.**

modRes permet effectuer l'analyse, **modRes_scrip** permet lancer toute une série d'analyses, **reduce** la réduction de données (nombre de résonances) et **mk_lisp_model** produit un fichier ascii à partir d'un fichier binaire.

L'analyse peut être utilisée sous UNIX, SGI ou Macintosh (dans Diphone).

Documentation on-line

Toutes les commandes décrites ci-dessous répondent à l'option –h en affichant à l'écran un court mode d'emploi, donnant la liste de tous les arguments possibles de la commande.

Formats des modèles et utilitaires de changement de format

Le format pour la sauvegarde des modèles de résonance peut être format binaire (toujours et par défaut), sdif (pour faire la resynthèse avec Fof ou filtres avec CHANT) et Ascii (deux formats, l'un avec toute l'information de les analyses et l'autre pour utiliser directement le fichier avec Max. Voir Edition d'un modèle, page 14). La commande **mk_lisp_model** produit un fichier ascii à partir d'un fichier binaire. Voici les options utilisables :

```
mk_lisp_model -h
Usage:mk_lisp_model -Les_flags_ci_dessous[valeur]
h :   pour mode d'emploi
i :   nom du fichier a convertir
d :   repertoire de lecture du modele binaire [courant]
o :   nom du fichier en sortie [-i.ll]
D :   repertoire d'ecriture du modele ascii [courant]
M :   ecriture du modele max [no ecriture]
NB: pas de blanc entre signe - et option
```

Préparation des sons

Les sons musicaux analysés doivent se rapprocher d'une « réponse impulsionnelle »: pizzicato (sans glissé de hauteur) pour une corde, percussion, slap… Pour les sons très courts, ceux qui comportent des attaques très riches, des effets de couplage, il pourra être difficile d'obtenir un résultat satisfaisant.

- Le son à analyser doit d'abord être numérisé avec un taux d'échantillonnage quelconque.

- Il peut ensuite être normalisé en amplitude.

- Il est recommendé de *fenêtrer les sons de telle façon que l'attaque soit calée au tout début du fichier de son*, sans être cependant amputée (utiliser un éditeur graphique, avec une résolution de l'ordre de l'échantillon).

- Le son doit être monophonique et d'une durée supérieure à 0.1 s.

ModRes, dans un environnement UNIX où SGI, utilise les fonctions de STtools pour ouvrir les fichiers de son. Cela veut dire qu'il est possible d'utiliser tous les formats acceptés par STtools (Ircam, AIFF, AIFC, etc.).

Conduite d'une analyse

Rappel du principe

L'analyse est interactive et itérative.

A chaque étape, deux fenêtres sont choisies sur le son. Sur les spectre estimés, les pics sont extraits, puis appariés. Lorsque l'appariement est possible, on considère que la paire de pics représente un mode de résonance du corps ayant produit le son analysé. De la position en fréquence et de l'évolution en amplitude des pics, on déduit le taux de décroissance). Cette première opération produit un modèle « élémentaire ».

Ce modèle est alors confronté au modèle obtenu lors de la précédente étape d'analyse. Les résonances modélisées sont appariées lorsqu'elles sont suffisamment proches en fréquence, et leur paramètres modifiés en fonction des anciennes et des nouvelles valeurs. Lorsque l'appariement n'est pas possible, les résonances sont conservées ou rajoutées sans modification. Cette deuxième opération produit un modèle « enrichi » qui constitue le produit final de l'étape d'analyse (fig. nº 1).

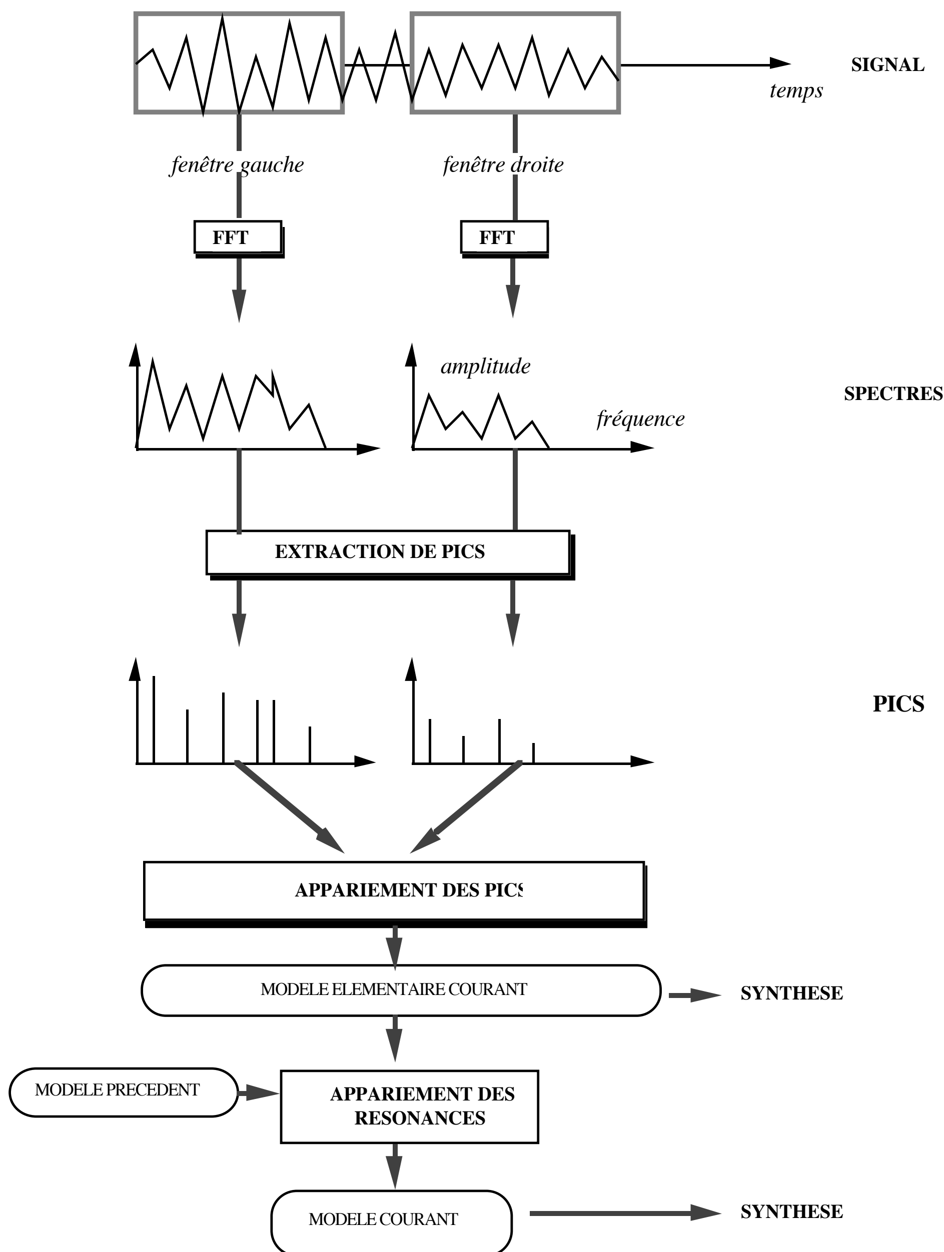


Figure 1 : Principe de l'analyse

L'enchaînement des étapes doit dans la version actuelle de l'analyse répondre au principe suivant :

l'analyse doit commencer avec deux petites fenêtres sur le son. Si le son à l'attaque au tout début du fichier c'est possible situer la première fenêtre avant le son (valeur négative pour le commencement de la fenêtre) pour essayer d'obtenir un bon analyse de l'attaque. C'est un mécanisme délicat d'utilisation; à chaque nouvelle étape, la taille des fenêtres et leur écartement doivent être augmentés, afin de gagner en résolution fré-

quentielle et d'affiner l'estimation du taux de décroissance des résonances. A chaque étape, l'utilisateur doit fournir un certain nombre de paramètres, et peut écouter le son correspondant à l'analyse élémentaire courante, et celui correspondant au résultat global (le modèle « enrichi ») courant. L'analyse est terminée, généralement au bout de 3 à 8 étapes, lorsque le résultat est jugé satisfaisant. A tout moment, il est possible de revenir en arrière en repartant d'une étape antérieure.

La figure nº 2 montre schématiquement tout le procès itérative. Après N étapes nous obtenons le modèle définitif mN. Dans la figure nº 3 nous avons la comparaison des spectres d'amplitude du son originel avec la resynthèse du premier modèle élémentaire (première fenêtre), avec la resynthèse du premier modèle enrichi (deuxième fenêtre) et finalement (troisième fenêtre) avec la resynthèse du cinquième modèle enrichi.

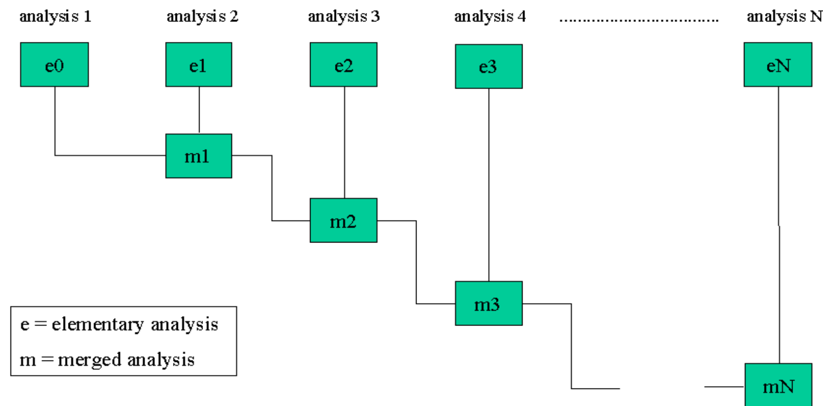


Figure nº 2 : Schème du procès itérative

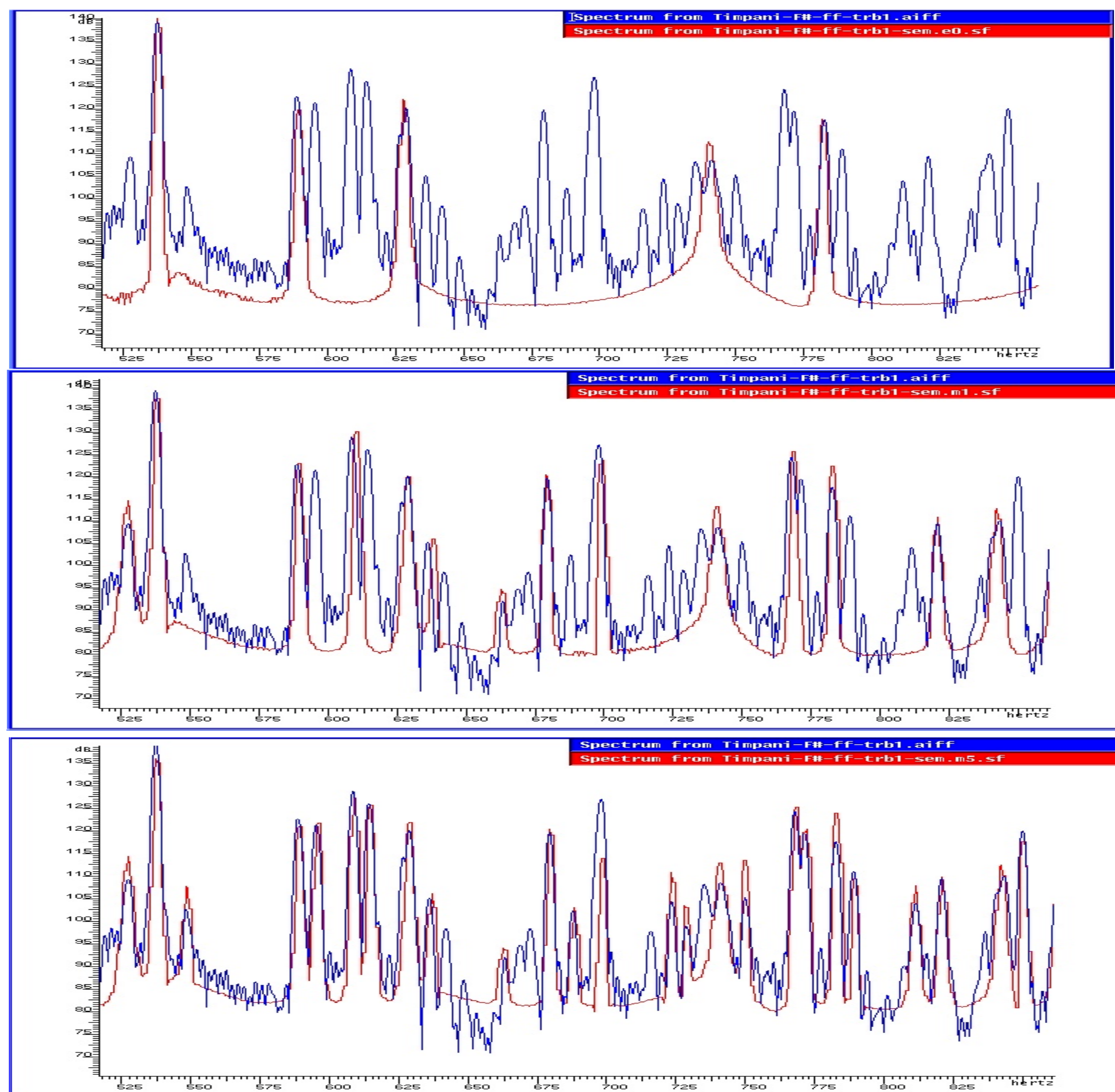


Figure nº 3 : Comparaison des spectres d'amplitude

Appel de la commande d’analyse

L'appel de la commande d'analyse avec l'option -h affiche à l'écran le message suivant, qui décrit brièvement les contrôles dont dispose l'utilisateur:

```
modRes -h

Usage:    modRes -Les_flags_ci_dessous [valeur]

controle des entrees sorties
            i :  nom du son a analyser
            o :  racine des noms de fichiers de sortie [racine d’entree]
            p :  modele precedent
            d :  repertoire de lecture ecriture des modeles [courant]
            D :  repertoire de lecture des fichiers de sons [courant]
            Z :  ecriture du modele sdif (n: no ecriture, e: ecriture du modele
                 elementaire, m: ecriture du modele enrichi, em ‘ou’ me: ecriture
                 du modele elementair et enrichi) [n]
            I :  ecriture du model ascii (n: no ecriture, e: ecriture du modele
                 elementaire, m: ecriture du modele enrichi, em ‘ou’ me: ecriture
                 du modele elementair et enrichi) [n]
            J :  ecriture du model max (n: no ecriture, e: ecriture du modele
                 elementair, m: ecriture du modele enrichi, em ‘ou’ me: ecriture du
                 modele elementair et enrichi) [n]
            v :  si positionne, commentaires affiches a l'ecran [non positionne]

contrôle pour la resynthese
            r :  duree [duration du son analyse]
            z :  option de resynthese (o: FOFs, i: Filtres) [o]

controle de l'estimation spectrale
            t :  temps pour la premiere fenetre [0.0 sec.]
            T :  temps pour la deuxieme fenetre [temps_premiere + taille de la
                 fenetre]
            w :  taille de la fenetre (en secondes) [0.1 sec.]
            c :  type de fenetre (0: Blackman exacte,
                 1: Blackman,
                 2: Blakman-Harris 3 (-67 dB),
                 3: Blackman-Harris 3 (-61 dB),
                 4: Blackman-Harris 4 (-92 dB),
                 5: Blackman-Harris 4 (-74 dB))       [4]
            f :  taille de la FFT (en secondes) [plus proche puissance de 2
                 superieure a la fenetre]

controle de l'extraction de pics
            P :  methode d'extraction de pics (0: all smaller,
                 1: one smaller,
                 2: mean smaller) [0]
            L :  largeur des pics (1 donne tous les maximums) [1]
            l :  hauteur des pics [2.00dB]
            a :  amplitude minimum des pics dans la premiere fenetre [-80.00dB]
            A :  amplitude minimum des pics dans la deuxieme fenetre [-96.00 dB]

controle de l'analyse elementaire (appariement des pics)
m :              mode d'appariement des pics (0: first match, 1: closest match) [1]
            s :  seuil en Hz. pour l'appariement [-y fois la resolution de
                 l'estimation spectrale]
            y :  facteur pour calcul de -s si non specifie [2.00]
```

```
controle de l'analyse merge (appariement des resonances)
            p :  modele a apparier au nouveau modele elementaire (si absent,
                 demarrage d'une nouvelle serie)
            M :  mode d'appariement des resonances (0: first match, 1: closest
                 match) [1]
            e :  flag mode doublement des resonances (0: off, 1: utiliser -b seul,
                 2: utiliser -b et -B) [0]
            V :  flag verification de la largeur de bande des resonances si non
                 retrouvees dans analyse courante [non positionne]
            S :  seuil en Hz. pour l'appariement [-Y fois la resolution de
                 l'estimation spectrale]
            Y :  facteur pour calcul de -S si non specifie [2.00]
            b :  valeur maximum de largeur de bande pour qu'une resonance puisse
                 etre doublee [12.00]
            B :  changement minimum en pourcentage de la largeur de bande par
                 rapport a l'analyse precedente pour qu'une resonance puisse etre
                 doublee [5.00]

NB : deux  flags obligatoires: -i (et -p pour continuer une analyse)
NB: pas de blanc entre flag et valeur
```

Entre crochets figurent les défauts utilisés lorsque l'option n'est pas spécifiée.

Contrôle des entrées-sorties

```
            i :  nom du son a analyser
            o :  racine des noms de fichiers de sortie [racine d'entree]
            p :  modele precedent
            d :  repertoire de lecture ecriture des modeles [courant]
            D :  repertoire de lecture des fichiers de sons [courant]
            Z :  ecriture du model sdif (n: no ecriture, e: ecriture du modele
                 elementaire, m: ecriture du modele enrichi, em 'ou' me: ecriture
                 du modele elementaire et enrichi) [n]
            I :  ecriture du model ascii (n: no ecriture, e: ecriture du modele
                 elementair, m: ecriture du modele enrichi, em 'ou' me: ecriture du
                 modele elementaire et enrichi) [n]
            J :  ecriture du model max (n: no ecriture, e: ecriture du modele
                 elementaire, m: ecriture du modele enrichi, em 'ou' me: ecrituredu
                 modele elementaire et enrichi) [n]
            v :  si positionne, commentaires affiches a l'ecran [non positionne]
```

Les noms des modèles produits par l'analyse sont formés par concaténation de la racine spécifiée par l'option -o (le nom du son par défaut), de « .e » pour un modèle « élémentaire », de « .m » pour un modèle « enrichi » (« merged »), puis du niveau d'analyse courant (la première étape correspondant au niveau 0). Ces fichiers ont format binaire. Les options –Z, -I et –J permettent les écrire en format Sdif (pour effectuer la resynthèse) ou Ascii (-I avec toute la information des étapes et –J pour la liste des résonances).

Lorsque l'option -p (le nom du modèle produit par l'étape précédente) n'est pas spécifiée, le niveau est 0 (on démarre une nouvelle analyse); dans le cas contraire, le niveau est incrémenté de 1 par rapport au niveau du modèle précédent.

L'option -v (« verbose »), affiche des commentaires sur le déroulement de l'analyse.

L'option –i es l'unique obligatoire. Les autres ont une valeur par défaut.

NB: *attention, le son doit avoir une durée supérieure ou égale à T+w (options précisées ci-dessous).* S'il n'est pas assez long, rajouter du silence pour qu'il atteigne la durée désirée.

Contrôle de la resynthèse

```
r :         dure spécifiée dans le modèle sdif [durée du son analysé]
z :         option de resynthèse (o: Fof, i: Filtres) [o]
```

modRes ne fait pas la resynthèse directement. Il écrit un fichier (avec les données spécifiées avec –r et –z) en format Sdif et CHANT, a partir de ce fichier il fait la resynthèse.

Contrôle de l'estimation spectrale

```
t :         temps pour la premiere fenetre [0.0 sec.]
T :         temps pour la deuxieme fenetre [temps_premiere + taille de la fenetre]
w :         taille de la fenetre (en secondes) [0.1 sec.]
            c :  type de fenetre;
                 0: Blackman exacte,
                 1: Blackman,
                 2: Blakman-Harris 3 (-67 dB),
                 3: Blackman-Harris 3 (-61 dB),
                 4: Blackman-Harris 4 (-92 dB),
                 5: Blackman-Harris 4 (-74 dB)) [4]
f :         taille de la FFT (en secondes) [plus proche puissance de 2 superieure a
                 la fenetre]
```

L'analyse utilise la transformée de Fourier rapide.

La fenêtre de type Blackman-Harris 4 normalement amène de meilleurs résultats.

Le temps de début de la première et de la deuxième fenêtre sur le son sont comptés par rapport au début du son. La première fenêtre peut avoir une valeur négative. Cela voudra dire qu' elle est située avant le début du son.

La taille de la fenêtre (la portion du son prise en compte pour l'estimation spectrale) peut être quelconque. La taille de la FFT peut être quelconque. L'utilisateur doit la spécifier en seconds mais, par défaut, c'est la plus proche puissance de 2 supérieure ou égale à la taille de la fenêtre et à la taille de la FFT pour le taux d'échantillonnage.

Contrôle de l'extraction de pics

```
P :         methode d'extraction de pics;
                 0: all smaller,
                 1: one smaller,
                 2: mean smaller) [0]
L :         largeur des pics (1 donne tous les maximums) [1]
l :         hauteur des pics [2.00dB]
a :         amplitude minimum des pics dans la premiere fenetre [-80.00dB]
A :         amplitude minimum des pics dans la deuxieme fenetre [-96.00 dB]
```

Trois variantes sont disponibles pour l'extraction de pics correspondant respectivement aux valeurs 0, 1 et 2 de l'option -P.

L'algorithme considère comme un pic tout maximum dans le spectre, dont l'amplitude est supérieure à -adB pour la première fenêtre, -AdB pour la seconde fenêtre, et qui soit entouré de part et d'autre par -L points d'amplitude inférieure.

Dans la première variante, les amplitudes de ces points (à l'exception des points immédiatement adjacent) doivent être inférieures de -l dB à l'amplitude du maximum.

Dans la deuxième variante, il suffit qu'un seul point (à l'exclusion des points immédiatement adjacent) soit inférieur de -l dB au maximum.

Dans la troisième variante, c'est la moyenne des amplitudes des points (y compris les points immédiatement adjacents) qui doit être inférieure de -l dB au maximum.

C'est clair que la première variante c'est la moins restrictive (plus résonances).

Pour les deux premières variantes, une valeur de 1 pour l'option -l revient à retenir tous les maximums supérieurs à -a ou -A dB.

Contrôle de l'appariement des pics

```
m :        mode d'appariement des pics (0: first match, 1: closest match) [1]
s :        seuil en Hz. pour l'appariement [-y fois la resolution de l'estimation
                spectrale]
y :        facteur pour calcul de -s si non specifie [2.00]
```

L'appariement des pics est effectué suivant un critère de proximité en fréquence: peuvent être appariés deux pics dont la distance dans le domaine des fréquences est inférieure à -s Hz. Si l'option -s n'est pas spécifiée, le seuil utilisé est égal à -y fois la résolution de l'estimation spectrale (1/fft_size en secondes).

Les deux jeux de pics sont parcourus par ordre de fréquences croissantes. Pour la valeur 0 de l'option -m, le premier appariement possible est retenu. Pour la valeur 1 de l'option -m, le meilleur appariement possible est recherché (i.e. ce sont les pics les plus proches qui sont appariés). Le mode 0 correspond au fonctionnement de la version Lisp de l'analyse, le mode 1 n'existe que dans la version C et doit être préféré (mais normalement donnent le même résultat).

Contrôle de l'appariement des résonances

```
p :        modele a apparier au nouveau modele elementaire (si absent,demarrage
                d'une nouvelle serie)
M :        mode d'appariement des resonances (0 : first match; 1 : closest match)
                [1]
e :        flag mode doublement des resonances (0: off,1: utiliser -b seul, 2:
                utiliser -b et -B) [0]
V :        flag verification de la largeur de bande des resonances si non
                retrouvees dans analyse courante [non positionne]
S :        seuil en Hz. pour l'appariement [-Y fois la resolution de l'estimation
                spectrale]
Y :        facteur pour calcul de -S si non specifie [2.00]
b :        valeur maximum de largeur de bande pour qu'une resonance puisse etre
                doublee [12.00]
B :        changement minimum en pourcentage de la largeur de bande par rapport a
                l'analyse precedente pour qu'une resonance puisse etre doublee
                [5.00]
```

Les modes d'appariement des résonances sont les mêmes que pour les pics. La différence c’est que maintenant les résonances sont conservées ou rajoutées sans modifications lorsque l’appariement n’est pas possible.

Deux options permettent d'activer des modes spéciaux:

- Lorsque l'option -V (vérification du taux de décroissance des résonances) est positionnée, le mécanisme suivant est activé: lorsqu'une résonance modèle produit par l'étape précédente n'est pas retrouvée dans l'analyse élémentaire de l'étape courante, on vérifie que le taux de décroissance de cette résonance est compatible avec la disparition observée. Si elle ne l'est pas, le taux de décroissance est augmenté dans nouveau modèle. Ce mécanisme est délicat d'utilisation (une résonance peut très bien échapper à une étape intermédiaire de l'analyse sans pour autant qu'elle soit « éteinte » dans le son), en particulier lorsqu'on utilise l'option -e.

- Lorsque l'option -e est positionnée, on cherche à modéliser les résonances dans le son dont la décroissance n'est pas exponentielle par deux filtres dont les taux de décroissance diffèrent, suivant le mécanisme suivant: si une résonance du modèle de l'étape précédente est retrouvée dans l'analyse élémentaire de l'étape courante avec un taux de décroissance significativement plus faible, la résonance est « doublée »; une première résonance, plus amortie, représente l'attaque, la deuxième, moins amortie mais d'amplitude plus faible, représente la deuxième partie de la réponse impulsionnelle. Pour que le « doublement » intervienne, deux conditions doivent être remplies: la largeur de bande de la résonance dans l'étape courante (qui mesure le taux d'amortissement) doit être inférieure à une valeur spécifiée par l'option -b, et le changement constaté par rapport à l'étape précédente doit être supérieure en pourcentage à la valeur spécifiée par l'option -B. Ce mécanisme est utile pour certains sons (piano par exemple), mais délicat d'utilisation.

Choix des paramètres et enchaînement des étapes successives

La première fenêtre d'analyse ne doit pas obligatoirement commencer au temps 0, mais en pratique, cette valeur de 0 amène de meilleurs résultats. C'est possible aussi commencer avant (temps <0). L'augmentation de la taille de la fenêtre d'analyse dépend de la durée et de la richesse de l'attaque du son à analyser. On pourra pour commencer doubler la durée à chaque étape. Si les temps de début des fenêtres n'est pas spécifié, la première débute toujours au temps 0, et la deuxième au temps correspondant à la fin de la première fenêtre. Important: le message *fin du son atteinte avant lecture des échantillons demandés* veut dire que le son est court pour les fenêtres spécifiées.

Exemple d'enchaînement

```
modRes -icloche -w.01
modRes -icloche -w.03  -pcloche.e0
modRes -icloche -w.05  -pcloche.m1
modRes -icloche -w.10  -pcloche.m2
```

C'est le contrôle auditif qui permet de bien mener les étapes, et la compréhension du principe d'analyse et une certaine expérience est nécessaire. Rappelons simplement ici que les petites fenêtres permettent d'extraire les résonances très amorties de l'attaque (généralement aiguës) et les résonances peu amorties les plus fortes, tandis que les grandes fenêtres permettent d'affiner l'estimation des paramètres de ces dernières et d'extraire les résonances peu amorties d'amplitude plus faible, ainsi que de séparer des résonances très proches responsables de battements. La taille maximum pour les fenêtres dépend de la mémoire disponible sur l'hôte. Les fenêtre de plus de 1 seconde ne sont utiles que pour des sons très résonants, pour séparer des résonances très proches.

Une fois l'analyse terminée, les fichiers correspondant aux modèles des étapes intermédiaires peuvent être supprimés; c'est l'un des derniers modèles « enrichis » (suffixe « .m ») qui constitue le modèle final.

Les modèles produits comportent une histoire des étapes de l'analyse; on pourra consulter dans la bibliothèque de modèles (Lisp ou C), sur les fichiers ascii, les paramètres utilisés pour les différents instruments. Cette histoire permet aussi de répéter une analyse, simplement en exécutant la liste des commandes fournies.

Si le résultat de la synthèse de l'analyse courante révèle des erreurs flagrantes, il est possible de revenir en arrière en précisant à l'aide de l'option -p le modèle « enrichi » de étape antérieure dont on souhaite repartir en variant les paramètres de l'analyse.

Lancement d'une analyse en mode non interactif

Le contrôle auditif de chacune des étapes est important pour réussir une analyse. Il est toutefois possible de lancer toute une série d'analyses avec la commande **modRes_script.** L'appel de la commande avec l'option –h affiche à l'écran le message suivant :

Usage:

```
modRes_script -Les_flags_ci_dessous [valeur]

i : nom du son a analyser. Flag obligatoire
d : repertoire d'ecriture des modeles [courant]
D : repertoire de lecture des fichiers de sons [courant]
o : racine des noms de fichiers de sortie [racine d'entree]
n : nom del script. Flag obligatoire
```

NB: blanc entre flag et valeur

Le script demandé a l'option –n est un autre fichier contenant toute la série d'analyses. L'utilisateur peut utiliser les scripts déjà implémentés (avec -nNom du fichier) ou peut construire son ensemble de commandes d'analyse (ou modifier les fichiers script déjà existants). Par exemple, avec la commande

```
modRes_script –iPiano –nMrscrip5fof
```

nous faisons toute une série d'analyses sur le son « Piano ». Les analyses viennent spécifiées dans Mrscrip5fof.

NB: nous avons déjà dit que le son doit avoir une durée supérieure ou égale a T (commencement de la deuxième fenêtre) + w (taille de la fenêtre). Peut être un script a été écrit pour analyser sons longs. Pourtant, si nous essayons d'analyser un son court nous aurons le message *fin du son atteinte avant lecture des echantillons demandés*. Il faut changer les durations des fenêtres du scrip (ou utiliser un autre).

Edition d'un modèle

Voici un exemple de modèle produit par l'analyse, en format Ascii (version -I)

```
*** INFORMATION DEL MODELE ***

numero total de resonances 282
model_type MODELE_MERGE
son Timpani
date Sun May 30 17:33:36 1999
sample_rate 44100.000000
niveau_analyse 5
```

```
niveau 0
  t1 0.000000 t2 0.097139 wsize 0.097139 fftsize 0.185760
  window type BLACKMAN_HARRIS_3_(-92_dB)  left_dB_floor -80.000000 right_dB_floor -
                96.000000 peak_width 1 peak_height 2.000000 peak_mode 0
  peak_match_type 1 peak_threshold 10.766602
  double_mode 0 max_d_band 0.000000 min_d_reldeltaT 0.000000 check_mode 0
                res_match_type 0 res_threshold 0.000000
  psize1 208 psize2 193 ele_model_size 103 model_size 103

niveau 1
  t1 0.000000 t2 0.194279 wsize 0.194279 fftsize 0.371519
  window type BLACKMAN_HARRIS_3_(-92_dB)  left_dB_floor -80.000000 right_dB_floor -
                96.000000 peak_width 1 peak_height 2.000000 peak_mode 0
  peak_match_type 1 peak_threshold 5.383301
  double_mode 0 max_d_band 12.000000 min_d_reldeltaT 5.000000 check_mode 1
                res_match_type 1 res_threshold 5.383301
  psize1 295 psize2 214 ele_model_size 150 model_size 213

niveau 2
  t1 0.000000 t2 0.388558 wsize 0.388558 fftsize 0.743039
  window type BLACKMAN_HARRIS_3_(-92_dB)  left_dB_floor -80.000000 right_dB_floor -
                96.000000 peak_width 1 peak_height 2.000000 peak_mode 0
  peak_match_type 1 peak_threshold 2.691650
  double_mode 0 max_d_band 12.000000 min_d_reldeltaT 5.000000 check_mode 1
                res_match_type 1 res_threshold 2.691650
  psize1 355 psize2 126 ele_model_size 89 model_size 267


modRes -iTimpani -R44100.000000 -t0.000000 -T0.097139 -w0.097139 -f0.185760 -P0 -L1
                -l2.000000 -a-80.000000 -A-96.000000 -m1 -s10.766602 -c4
modRes -iTimpani -R44100.000000 -t0.000000 -T0.194279 -w0.194279 -f0.371519 -P0 -L1
                -l2.000000 -a-80.000000 -A-96.000000 -m1 -s5.383301 -c4 -
                pTimpani.e0 -M1 -e0 -V -S5.383301 -b12.000000 -B5.000000
modRes -iTimpani -R44100.000000 -t0.000000 -T0.388558 -w0.388558 -f0.743039 -P0 -L1
                -l2.000000 -a-80.000000 -A-96.000000 -m1 -s2.691650 -c4 -
                pTimpani.m1 -M1 -e0 -V -S2.691650 -b12.000000 -B5.000000


model_size 267
   1        18.500       -67.510 (0.000)         0.873    0    0
   2        27.179       -69.332 (0.000)         0.618    0    0
   3        32.656       -39.637 (0.010)         0.602    0    0
   4        38.020       -59.071 (0.001)         0.931    0    0
   5        55.647       -67.359 (0.000)          1.76    0    0
...
```

On trouve d'abord le numéro de résonances et le type du modèle (« élémentaire » (MODELE_ELE) ou « enrichi » (MODELE_MERGE)), et des informations générales, puis les valeurs des options utilisées à chacun des niveaux d'analyse, puis la liste des commandes permettant de reproduire l'analyse à partir du son (les noms des répertoires d'entrée et de sortie des sons et des modèles ne sont pas conservés).

On trouve ensuite le nombre des résonances du modèle suivi de la liste des paramètres des résonances: numéro de la résonance, fréquence, puissance en dB (puissance linéaire entre parenthèses), largeur de bande. Les deux valeurs suivantes indiquent si la résonance a été doublée suivant le mécanisme décrit plus haut (0 pour une résonance simple, 1 pour une résonance qui a été doublée, -1 pour la résonance « double » qui lui correspond), et si son taux d'amortissement a été vérifié (1 pour vérifiée, 0 pour non vérifiée).

Avec l'option –J c'est possible construire un autre fichier ascii avec différent format. Ici nous avons le numéro de la résonance, fréquence, puissance linéaire et largeur de bande.

```
0, 18.499630 0.000421 0.872843;
1, 27.178673 0.000341 0.618150;
2, 32.655804 0.010426 0.601608;
3, 38.019592 0.001113 0.930745;
4, 55.647129 0.000429 1.757132;
5, 63.312294 0.001380 1.305699;
6, 75.996735 0.000535 2.844953;
7, 86.133293 0.000407 2.855310;
8, 127.917404 0.002997 1.717828;
9, 185.600052 0.131805 1.823762;
10, 234.544113 0.000194 1.971352;
11, 262.347626 0.000443 1.517185;
12, 274.224121 0.013323 0.897565;
```

...

Réduction de données

Une réduction de données peut être appliqué aux modèles produits par l'analyse, *en bout de chaîne* (un modèle « réduit » ne peut pas être utilisé pour reprendre une analyse).

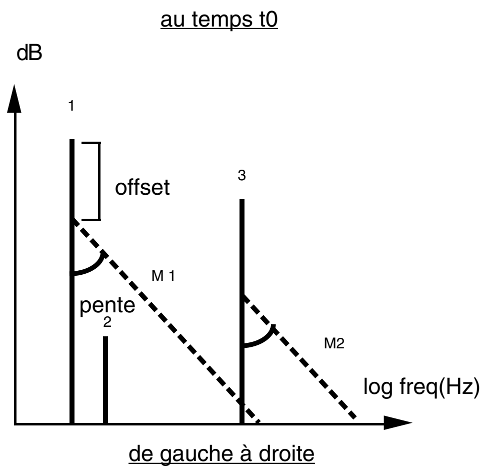
Principe

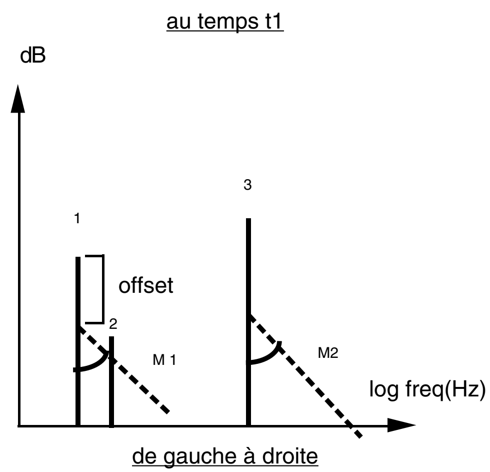
La réduction de donnée prend en compte de façon simplifiée les effets de masquage entre résonances. Les caractéristiques du masque sont définies par l'utilisateur, ce qui permet de réduire plus ou moins radicalement les données. L'algorithme de réduction peut être appliqué à différents instants de la réponse (simulée) du modèle à une excitation de type dirac; seules sont alors supprimées les résonances qui sont masquées à chacun des instants choisis. Il peut aussi être appliqué sur l'ensemble de la réponse impulsionnelle (cette variante n'existe pas dans la version Lisp).

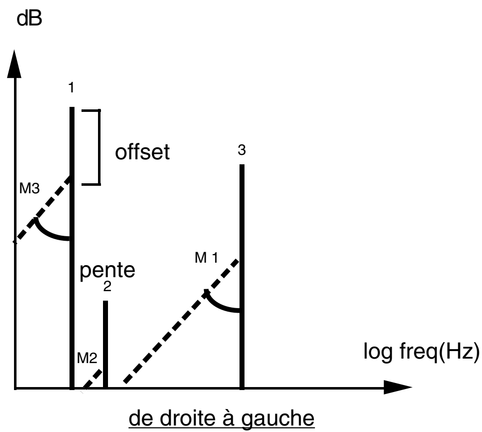
Le masque est défini par un « offset » (en dB.) et par une pente (en dB. par octave); pour chaque résonance, un masque est calculé, qu'on peut visualiser dans une représentation log-fréquence/dB-amplitude du modèle comme une demi-droite partant du point défini par la fréquence de la résonance et par son amplitude en dB diminué de « offset » dB, et dont la pente est exprimée en dB par octave. Toute résonance qui « émerge » des masques calculés, c'est -à-dire toute résonance dont l'amplitude calculée à l'instant choisi est supérieure à la valeur du masque pour la fréquence de la résonance, est retenue. Les autres sont écartées. L'algorithme prend en compte l'effet de masquage de l'aigu par la grave, et de la grave par l'aigu (cf. figure no 1). Il ne prend effet qu'au-dessus d'un niveau plancher (l'option -P de la commande), c'est-à-dire qu'une résonance masquée en dessous de -(-P) dB est conservée.

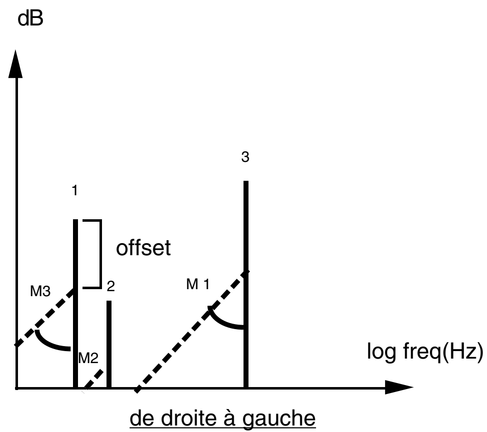
Le modèle simplifié est écrit sur le répertoire choisi, et le son correspondant à la resynthèse est calculé, le contrôle auditif du résultat étant essentiel.

Les résonances peuvent aussi êtres éliminés suivant des critères absolus (fréquences et largeurs de bandes minimales et maximale, amplitude minimale).









offset en Hz.
pente en dB/octave
Les résonances sont numérotées
par ordre de fréquences croissantes,
les masques M par ordre d'utilisation

Figure. Principe de la réduction de données

Appel de la commande de réduction

L'appel de la commande de réduction de données avec l'option -h affiche à l'écran le message suivant, qui décrit brièvement les contrôles dont dispose l'utilisateur:

```
Usage :

reduce -Les_flags_ci_dessous [valeur]
i : nom du fichier a reduire
o : nom du fichier de sortie [entree.r]
d : repertoire de lecture des modeles [courant]
D : repertoire d'ecriture des modeles [courant]
I : ecriture du modele ascii [no ecriture]
J : ecriture du modele max [no ecriture]
Z : ecriture du modele sdif (n: no ecriture, o: FOFs, i: Filtres) [n]
r : duree especifie dans le modele sdif [2.0]
M : flag de reduction par effet de masquage
                 O : offset du masque en dB [20.00]
S : pente du masque en dB/octave [34.00]
P : amplitude minimum pour la reduction en dB [70.00]
t : flag pour mode liste de temps (calcul du masque
    pour chacun des temps specifies)[off]
l : un temps en secondes a rajouter a liste pour option -t
                 T : mode continu: calcul du masque sur toute la reponse
     impulsionnelle a partir du temps specifie [on]
w : un temps en secondes pour flag -T [0.00]
A : flag de reduction suivant criteres absolus
f : frequence minimum (Hz) [20.00]
F : frequence maximum (Hz) [20000.00]
a : amplitude minimum (dB) [70.00]
b : largeur de bande minimum (Hz) [0.00]
B : largeur de bande maximum (Hz) [20000.00]
v : si positionne, commentaires affiches a l'ecran
NB: -t et -T sont exclusifs
NB: pas de blanc entre signe - et option
```

Les options -M et -A peuvent être utilisées simultanément.

Voici un exemple d'appel de la commande de réduction:

```
reduction -i:vib-G2 -M  -T
```

Voici un fichier ascii correspondant à une modèle « réduit »:

```
model_type MODELE_REDUIT
original vib-G2
date 28/11/90_14:41:29
abson: 1
 min_freq 20.000000
 max_freq 20000.000000
 min_ampl -70.000000
 min_band 0.000000
 max_band 20000.000000
maskon: 1
 slope 34.000000
 offset 20.000000
```

```
mask_floor -70.000000
timelistmode: 0
          ntimes 1
          timelist 0.000000
          wt 0.000000

reduire -ivib-G2 -A -f20.000000 -F20000.000000 -a-70.000000 -b0.000000 -
                B20000.000000 -M -S34.000000 -O20.000000 -P-70.000000 -T -
                w0.000000

model_size 40
   1        27.690    -65.437    0.32    0    0
   2        29.806    -66.388    0.362   0    0
   3        34.523    -49.554    0.195   0    0
   4        41.478    -67.527    0.696   0    0
   5        47.591    -43.851    0.182   0    0
   6        71.516    -68.833    0.551   0    0
   7        90.920    -60.949    1.43    0    0
   8       171.160    -49.098    3.39    0    0
   9       196.652    -24.250    0.138   0    0
  10       393.352    -46.459    0.605   0    0
  11       499.694    -68.503    2.73    0    0
  12       508.655    -48.197    7.9     0    0
  13       552.669    -69.970    2.31    0    0
  14       587.700    -64.415    0.344   0    0
  15       589.915    -47.352    0.594   0    0
  16       786.623    -35.179    0.204   0    0
  17       888.321    -51.473    5.65    0    0
  18       895.916    -45.162    12.4    0    0
  19       967.814    -58.843    1.7     0    0
  20       983.133    -61.016    0.254   0    0
  21      1581.176    -57.436    0.471   0    0
  22      2003.512    -29.321    1.27    0    0
  23      2006.710     -9.337    1.21    0    0
  24      3046.802    -31.721    0.722   0    0
  25      3050.696    -33.223    3.83    0    0
  26      3619.364    -49.355    2.51    0    0
  27      3623.105    -21.167    4.88    0    0
  28      4015.902    -42.213    4.15    0    0
  29      4907.378    -24.064    6.26    0    0
  30      5686.889    -53.074    5.24    0    0
  31      5723.582    -45.190    5.24    0    0
  32      6023.894    -57.616    3.42    0    0
  33      6820.997    -56.957    7.03    0    0
  34      6925.660    -42.997    6.47    0    0
  35      8150.777    -64.919    4.87    0    0
  36      8213.490    -67.177    3.52    0    0
  37      8938.438    -52.730    8.48    0    0
  38      8990.316    -44.179    6.85    0    0
  39     10119.490    -66.333    7.13    0    0
  40     10242.320    -63.332    3.49    0    0
```

Utilisation

Un offset de 20 dB. et une pente de 34 dB. par octave donnent généralement un résultat qui n'altère pas de façon significative la qualité de la resynthèse. Un offset et une pente plus petits déterminent une réduction de donnée plus radicale.

Avec l'option -t, le premier temps fourni doit généralement correspondre au début de la réponse (temps 0.). Les résonances de l'attaque sont alors conservées, mais les résonances plus faibles qui constituent l'essentiel de la réponse au-delà de l'attaque disparaissent. C'est pourquoi il est indiqué de fournir des arguments optionnels pour d'autres temps de réduction (chacun de ces temps étant spécifiés à l'aide de l'option -l), afin de préserver ces résonances, ou d'effectuer la réduction de données sur l'ensemble de la réponse impulsionnelle (option -T). Avec cette dernière option, les effets de masquage ne sont calculés qu'à partir du temps spécifié par l'option -w de la réponse impulsionnelle du modèle.

Références

Le principe général de l'obtention et de l'utilisation de ces modèles a été décrit ailleurs; pour en savoir plus, on peut se reporter aux documents suivants (et notamment aux documents 4 et 6):

(1) P-F. Baisnée, J-B. Barrière, O. Koechlin, M. Puckette & R. Rowe (1986) Real-time interaction between musicians and computer: performance utilizations of the 4X, *Proceedings of 1986 International Computer Music Conference, La Haye*, Berkeley: Computer Music Association, p.237-240.

(2) J-B.Barrière, Y. Potard & P-F. Baisnée (1985) Models of Continuity between Synthesis and Processing for the Elaboration and Control of Timbre Structures, *Proceedings of 1985 International Computer Music Conference,Vancouver*, Berkeley: Computer Music Association, p.193-198.

(3) Y. Potard, P-F. Baisnée & J-B.Barrière (1986) Experimenting with Models of Resonance Produced by a New Technique for the Analysis of Impulsive Sounds, *Proceedings of 1986 International Computer Music Conference, La Haye*, Berkeley: Computer Music Association, p.269-274.

(4) Yves Potard, P-F. Baisnée, J-B. Barrière (1986) Interaction matériau organisation - version préliminaire d'un rapport de recherche, *document interne Ircam.*

(5) Yves Potard, P-F. Baisnée, J-B. Barrière (1988) Méthodologie de synthèse du timbre: L'exemple des modèles de résonance, *Paris, Christian Bourgois.*

(6) Baisnée P-F. Mode d'emploi des outils développés à l'Ircam

(7) Baisnée P-F. Mode d'emploi de la version Csound

(8) Baisnée, Barrière, Baudot, Freed ....(1989) Texte ICMC 89

(9) Baisnée P-F. Obtention de modèles de résonance (version en C). Ircam 1990.

Le texte (4) décrit le principe de l'obtention et de l'utilisation des modèles.

Le texte (6) décrit l'utilisation des outils disponibles pour l'analyse et la synthèse dans l'environnement CHANT/FORMES.

Index